

Contribution à l'étude de la focalisation prosodique en français¹

Rémi Godement-Berline

Univ. Paris Diderot, Sorbonne Paris Cité

LLF (UMR 7110), 8 place Paul Ricœur, 75013 Paris, France

remi.godement@linguist.univ-paris-diderot.fr

RESUME

Cette étude porte sur la focalisation prosodique en français dans plusieurs styles de parole (parole spontanée et lecture ou interprétation par des acteurs). Nous attribuons à la focalisation des fonctions sémantico-pragmatiques ou emphatiques. Un groupe de dix experts en prosodie a relevé les occurrences de focalisation dans le corpus d'étude. Les résultats confirment que la focalisation est réalisée par une augmentation de hauteur et de durée. Ils diffèrent de la littérature précédente du point de vue du type de contour prosodique employé sur les occurrences de focalisation et de la présence d'accent initial. Des problèmes méthodologiques sont soulevés concernant l'analyse des contours terminaux et de la désaccentuation.

ABSTRACT

Contribution to the study of prosodic highlighting in French.

This paper studies prosodic highlighting in French in several types of speech (spontaneous speech and reading aloud or performing by actors). We take it that prosodic highlighting fulfills semantic-pragmatic or expressive functions. A group of ten prosody experts annotated the occurrences of prosodic highlighting in the corpus of study. Results confirm that prosodic highlighting is realized through an increase in pitch and duration. They differ from previous studies concerning the type of prosodic contour on occurrences of prosodic highlighting and the presence of initial secondary accent. Methodological issues are raised concerning the analysis of terminal contours and deaccenting.

MOTS-CLES : prosodie, focalisation, accentuation, contour, parole spontanée

KEYWORDS: prosody, prosodic highlighting, accent, contour, spontaneous speech

1 Introduction

1.1 Présentation de l'étude

Cette étude porte sur la réalisation phonétique et la caractérisation phonologique du phénomène de la focalisation prosodique en français. La définition de la focalisation que nous adoptons est la suivante : un soulignement prosodique de constituant (par divers moyens tels que l'accentuation ou les variations de registre et de tempo) remplissant une fonction soit sémantico-pragmatique soit emphatique. Le premier type de fonction correspond au marquage de focus (informationnel,

¹ Une partie des travaux présentés dans cette étude a déjà fait l'objet d'une publication en anglais (Godement-Berline, à paraître).

contrastif, quantificationnel/associatif, verum), et le second type correspond aux fonctions d'insistance et d'expressivité (e.g. Rossi, 1999 ; Di Cristo, 1999). Deux exemples sont donnés ci-dessous, tirés de notre corpus annoté par les experts (Fig. 1, 2).

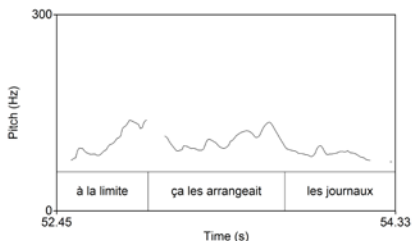


FIGURE 1 : Focalisation prosodique sur *arrangeait* (fonction : marquage de focus).

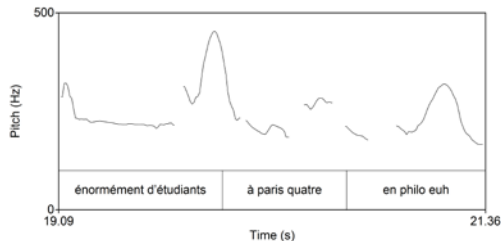


FIGURE 2 : Focalisation prosodique sur *énormément d'étudiants* (fonction : insistance).

De nombreuses études similaires ont déjà été menées (e.g. Di Cristo, 1999 ; Rossi, 1999 ; Jun et Fougeron, 2000 ; Astésano, 2001 ; Martin, 2009 ; De Looze, 2010 ; Delais-Roussarie & Di Cristo, à paraître). Cette étude se distingue par une définition fonctionnelle plus précise de la focalisation prosodique (qui est souvent limitée soit au marquage de focus soit à l'emphase) et par une méthodologie empruntée à la linguistique expérimentale faisant appel à un protocole d'évaluation explicite. Cependant, cette étude a pour inconvénient d'utiliser un corpus comportant plusieurs styles de parole différents ; ceci est dû au fait que l'expérience décrite ci-dessous s'inscrit dans un programme plus large consistant à étudier les différences de réalisation de la focalisation prosodique selon le style de parole (cf Godement-Berline, à paraître).

1.2 Etat de l'art

Un des principaux traits de la focalisation décrit par la littérature est une forte augmentation de F0, de durée et d'intensité, bien que les études précédentes ne soient pas unanimes concernant les deux derniers paramètres (e.g. Astésano, 2001 ; De Looze, 2010 ; Rossi, 1999). Un deuxième trait souvent décrit est la présence d'un accent initial sur le constituant focalisé, réalisé selon Delais-Roussarie & Di Cristo (à paraître) par une « augmentation significative de la durée de la consonne », une « glottalisation initiale si la première syllabe du mot lexical ne commence pas par une consonne » et la « présence d'une variation de hauteur perçue comme un pic mélodique extra haut ».

La présence d'un contour terminal sur le constituant focalisé est également observée, surtout concernant le marquage de focus. La définition de ce trait est toutefois problématique. Existe-t-il une forme phonologique objective de la finalité, ou bien s'agit-il d'un phénomène purement perceptif, ou qui interagit avec d'autres niveaux linguistiques (en particulier la syntaxe et la pragmatique) ? La première position est défendue notamment par Martin (2009) ou Rossi (1999), selon qui le contour terminal atteint le haut ou le bas du registre du locuteur. Cependant, certaines études (e.g. Post 2002)

montrent que des contours perçus comme terminaux peuvent ne pas correspondre à cette forme et, inversement, que des contours correspondant à cette forme peuvent ne pas être perçus comme terminaux.

Enfin, de nombreuses études sur la focalisation prosodique décrivent la présence d'une « désaccentuation » avant ou après le constituant focalisé, un trait lui aussi difficile à caractériser. Le cas le plus étudié, celui de la désaccentuation post-focale dans un énoncé d'assertion, est souvent décrit comme un contour « plat » dans le niveau bas du registre du locuteur (e.g. Rossi, 1999 ; Jun & Fougeron, 2000), éventuellement accompagné d'une accélération du débit d'articulation et d'une chute de l'intensité (De Looze 2010 ; Rossi 1999). Cette compression de la hauteur n'entraîne toutefois pas de suppression des groupements prosodiques (*dephrasing*), les frontières finales de groupe étant marquées par des allongements (e.g. Jun & Fougeron, 2000). La désaccentuation post-focale dans des énoncés à modalité interrogative avec contour terminal montant est décrite soit comme un contour montant (Rossi, 1999), soit comme un plateau de niveau haut ou moyen-haut suivi d'une montée (Jun et Fougeron, 2000), le ton final étant généralement lui aussi terminal (H%).

2 Méthodologie

2.1 Elicitation des données

Le corpus est constitué de deux extraits de parole spontanée et de huit enregistrements de locuteurs lisant à haute voix ou « interprétant » de mémoire la transcription de ces extraits spontanés (protocole RepTask ; cf Laurens et al., 2011). Le premier extrait spontané provient du corpus CID (Bertrand et al., 2008) et le second a été enregistré pour les besoins de l'expérience. Il s'agit dans les deux cas de conversations de type « bavardage » entre deux locuteurs ; dans les extraits retenus, un des deux locuteurs raconte une anecdote drôle et insolite, avec peu d'interruption de la part de l'autre. Les locuteurs sont deux hommes de la région Provence-Alpes-Côte d'Azur dans le premier extrait et deux femmes de Paris dans le second, tous âgés entre 25 et 45 ans et chercheurs ou doctorant(e)s en linguistique. Les locuteurs des versions lues et interprétées sont deux femmes et deux hommes de Paris ayant entre 20 et 40 ans et possédant tous, à des degrés divers, une expérience dans le jeu d'acteur. Leur tâche a consisté à restituer le texte comme s'ils étaient le locuteur d'origine et participaient réellement à la conversation.

Les enregistrements spontanés ont été réalisés dans une pièce calme avec un minimum de bruit de fond, à l'aide d'un micro-casque dans le premier cas et d'un micro portable Zoom H2 (format 16 bit/44.1 kHz, WAV) dans le second. Les enregistrements lus et interprétés ont été réalisés dans une chambre sourde à l'Université Paris Diderot à l'aide d'un microphone studio Rode NT1-A, d'une interface audio Roland Quad-Capture et du logiciel Audacity (format 16 bit/44.1 kHz, WAV). Le corpus (accompagné de fichiers d'alignement au format Textgrid) est disponible à l'adresse <http://www.lif.cnrs.fr/reptask>.

2.2 Annotation des occurrences de focalisation prosodique

Un groupe de dix experts en prosodie a annoté les occurrences de focalisation prosodique dans le corpus, en prenant pour unité d'annotation le mot. Les experts sont des chercheurs, doctorant(e)s ou

étudiant(e)s de niveau Master en phonétique, phonologie ou pragmatique ayant de l'expérience dans l'étude de la prosodie. Chaque expert a annoté quatre enregistrements, de sorte que chaque enregistrement a été annoté par quatre experts. Les experts ont écouté les enregistrements (sans limite de temps et sans l'aide d'un logiciel d'analyse acoustique permettant de visualiser le signal et ses caractéristiques prosodiques) et se sont référés pour l'annotation aux traits de la focalisation prosodiques cités plus haut (seuls ou en combinaison) : augmentation de hauteur, de durée ou d'intensité, présence d'accent initial, contour terminal sur le constituant focalisé, et désaccentuation avant ou après le constituant focalisé. Il a été demandé aux experts de ne pas annoter les accents uniquement rythmiques (marquant la frontière initiale ou finale d'un groupe prosodique). Seuls les mots annotés par au moins trois sur quatre experts ont été considérés comme des occurrences de focalisation prosodique.

2.3 Analyse prosodique

Les enregistrements ont été segmentés en mots, syllabes et phones à l'aide de l'extension de Praat EasyAlign (Goldman, 2011). Ils ont ensuite été analysés à l'aide de l'extension de Praat Prosogram (Mertens, 2004), qui donne une série de mesures pour chaque syllabe (en utilisant la segmentation préalable). Basé sur un modèle psycho-acoustique de la perception tonale, Prosogram mesure la F0 sur les portions voisées significatives uniquement, déterminées ici par le programme au sein de la rime syllabique. La F0 a été convertie en demi-tons (relatifs à 1 Hz) afin de pouvoir permettre la comparaison entre locuteurs. La durée a été normalisée par rapport à la structure syllabique en divisant la durée de chaque syllabe par le nombre de phones dans la syllabe, et par rapport au débit de parole des locuteurs en convertissant les valeurs précédentes en z-scores pour chaque locuteur.

Chaque occurrence de focalisation prosodique a été analysée visuellement et auditivement sur Praat par l'auteur afin d'en caractériser le contour prosodique, en utilisant la transcription ToBI pour le français (cf Delais-Roussarie et al., 2015). L'empan du contour au niveau du mot (i.e., sur quelle(s) syllabe(s) du constituant focalisé s'étend le contour) a également été déterminé par l'auteur. Enfin, la présence d'un accent initial sur la frontière gauche du constituant focalisé a été déterminée à l'aide de la fonction de détection automatique de proéminences du logiciel Analor (Avanzi et al., 2008), à partir de la segmentation syllabique préalable.

3 Résultats

Le corpus est composé d'un total de 5644 mots et de 7800 syllabes. Il contient 54% de mots lexicaux. Il contient 56,6% de syllabes de type CV, 14,8% de type V, 13,6% de type CVC, 9,8% de type CCV et 5,2% d'un autre type.

Le taux d'accord entre les annotations des experts a été obtenu en calculant le Kappa de Fleiss pour chaque enregistrement et en faisant la moyenne pour le corpus entier et pour chaque style de parole. Le taux d'accord pour le corpus entier est légèrement significatif (0,273, $z = 18,79$, $p = 0$). On observe des différences entre les styles de parole (taux d'accord plus haut pour la parole spontanée, suivie de l'interprétation et de la lecture).

La fréquence d'occurrence de focalisations prosodiques dans le corpus est faible (11,22 %). Elle est plus haute pour l'interprétation (14,11 %), suivie de la lecture (10,59 %) et de la parole spontanée (6,59 %) (les différences sont significatives en utilisant des intervalles de confiance à 95%).

3.1 F0 et durée syllabique moyennes

On observe une forte différence entre la F0 moyenne et la durée moyenne des syllabes comportant une focalisation et celles des autres syllabes du corpus (Fig. 3). Les données ont été analysées au moyen d'un modèle linéaire mixte, avec la présence de focalisation comme effet fixe et le locuteur, le groupe d'experts en prosodie et le style de parole comme effets aléatoires. Les p-values ont été obtenues au moyen d'un test du rapport des vraisemblances entre le modèle complet et un modèle sans l'effet fixe. Les différences sont significatives à la fois pour la F0 ($\chi^2(1) = 532,8$, $p < 0,01$) et pour la durée ($\chi^2(1) = 71,8$, $p < 0,01$).

Les différences de F0 et de durée, entre les styles de parole, des syllabes comportant une focalisation ont également été testées au moyen d'un modèle linéaire mixte avec le style de parole comme effet fixe et le locuteur et le groupe d'experts comme effet aléatoires, en utilisant à nouveau une réduction du modèle et un test du rapport des vraisemblances pour obtenir les p-values. Aucune différence ne s'avère être significative, à la fois pour la F0 ($\chi^2(2) = 0,137$, $p > 0,05$) et pour la durée ($\chi^2(2) = 3,617$, $p > 0,05$).

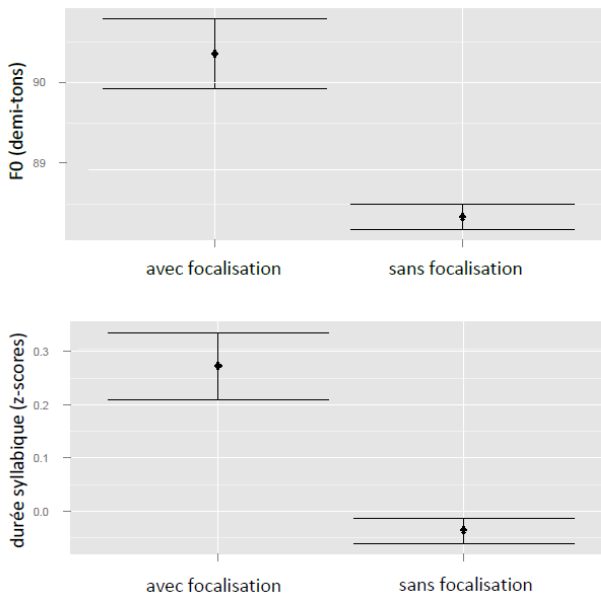


FIGURE 1: F0 et durée syllabique moyennes des syllabes comportant une focalisation prosodique et des autres syllabes du corpus

3.2 Caractérisation phonologique

On observe des différences de fréquence dans le corpus concernant le type de contour employé sur les occurrences de focalisation, le type de frontière de groupe prosodique réalisée, l'empan du contour au niveau du mot et la présence d'un accent initial. On n'observe pas de différences significatives selon ces mêmes variables entre les trois styles de parole. Ceci est confirmé par des tests du χ^2 de Pearson pour le type de contour ($\chi^2(12) = 13,20, p > 0,05$), le type de frontière ($\chi^2(4) = 8,78, p > 0,05$) et l'empan du contour ($\chi^2(6) = 7,78, p > 0,05$). Pour la présence d'accent initial, un modèle linéaire mixte a été employé avec le style de parole comme effet fixe et le locuteur et le groupe d'experts comme effets aléatoires, et n'a pas révélé de différence significative ($\chi^2(2) = 4,07, p > 0,05$).

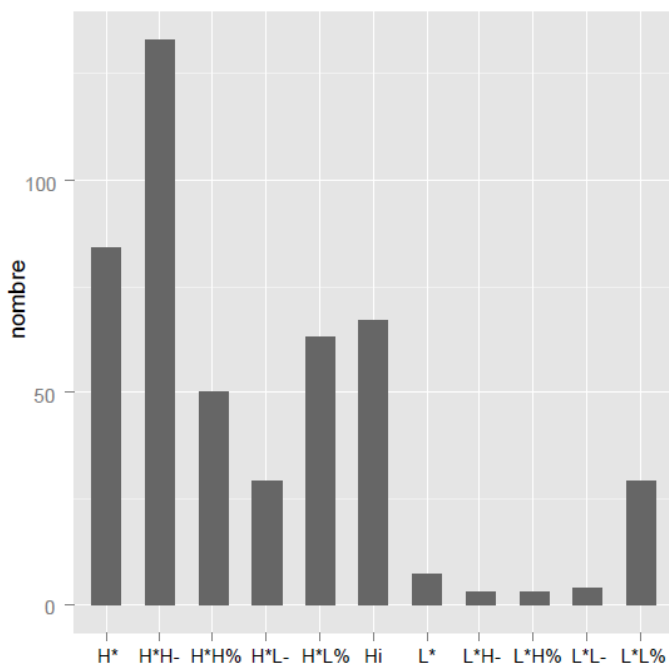


FIGURE 2 : Contours prosodiques employés sur les occurrences de focalisation

La majorité des contours sont montants (Fig. 4). Le contour le plus fréquent est H*H- (28,2 %), un contour de continuation majeure à la fin d'un groupe intermédiaire (ip), suivi de H* (17,8 %) et de Hi (14,2 %), un accent initial au début d'un groupe de mots focalisé. Les deux types de groupe prosodique les plus fréquemment réalisés à la frontière (initiale ou finale) des occurrences de focalisation sont le groupe intermédiaire (35,8 %) et le groupe accentuel (33,5 %). Lorsque le constituant focalisé fait plus d'une syllabe, les contours sont majoritairement réalisés sur la dernière syllabe (72,7 %). Une minorité de constituants focalisés plurisyllabiques comporte un accent initial (41,2 %).

4 Discussion

Les résultats confirment les études précédentes sur certains points et les contredisent sur d'autres. Les occurrences de focalisation prosodique présentent une nette augmentation de F0 et de durée syllabique, comme le notent Di Cristo (1999), Rossi (1999), Jun et Fougeron (2000), De Looze (2010) et Astésano (2001). Nous n'avons pas pu faire de mesures d'intensité car la distance micro-locuteur n'est pas fixe dans nos enregistrements (ce afin de ne pas gêner la spontanéité et l'expressivité des locuteurs). En revanche, notre caractérisation phonologique est différente de celle des mêmes auteurs, notamment car le contour prosodique le plus fréquent sur les occurrences de focalisation est le contour de continuation H*H- et car on observe une présence relativement faible d'accent initial. Le fait qu'on observe de façon minoritaire des frontières de groupe intonatif est contradictoire avec la fonction de marquage de focus (qui appelle normalement la réalisation d'un contour terminal, cf Di Cristo, 1999 ; Rossi, 1999 ; Martin, 2009) mais pas nécessairement avec les fonctions d'insistance et d'expressivité. En effet, l'insistance est décrite comme marquée principalement par la présence d'un accent initial (Di Cristo, 1999 ; Delais-Roussarie & Di Cristo, à paraître ; Astésano, 2001) ; le marquage de l'expressivité est quant à lui amalgamé avec le contour terminal (Rossi, 1999 ; Di Cristo, 1999) ou bien affecte l'énoncé entier (Di Cristo, 1999).

Plusieurs autres analyses de la focalisation pourraient être menées à partir du corpus, mais certaines présentent de sérieux obstacles méthodologiques. Comment déterminer, par exemple, la présence d'un contour terminal sur le constituant focalisé puisque, comme le montrent des études telles que Post (2002), ce contour ne possède aucune forme phonologique stable ? De même, comment objectiver la présence de désaccentuation sur la séquence précédant ou suivant le constituant focalisé ? Outre la variation de réalisation observée, il est difficile de différencier la désaccentuation de manière quantitative d'une simple baisse de hauteur (et de durée et d'intensité). Une solution possible serait de faire des mesures de différences de registre et de tempo (à l'aide par exemple des extensions de Praat ADoReVA et ADoTeVA ; cf De Looze, 2010) sur de la parole de laboratoire, par exemple sur des énoncés assertifs présentant un marquage de focus informationnel étroit, afin d'obtenir des seuils de différence à utiliser ensuite pour déterminer la présence de désaccentuation dans d'autres types de parole.

Cette étude sera prochainement complétée par une analyse des configurations tonales de la focalisation et de sa réalisation au niveau infrasyllabique (cf Astésano, 2001). Nous étudierons également l'influence de la catégorie lexicale des constituants focalisés.

5 Conclusion

Cette étude porte sur la focalisation prosodique en français, dans ses fonctions aussi bien sémantico-pragmatiques qu'emphatiques. Elle emploie une méthodologie expérimentale faisant appel à un protocole d'évaluation explicite. Nos résultats confirment que la focalisation est réalisée par une augmentation de hauteur et de durée. Ils diffèrent de la littérature précédente du point de vue de l'interprétation phonologique de la focalisation, notamment le type de contour prosodique employé et la présence d'accent initial. Une prochaine étude portera sur la réalisation prosodique des différentes fonctions de la focalisation (marquage de focus, insistance, expressivité).

Remerciements

Ce travail a bénéficié d'une aide de l'Etat gérée par l'Agence Nationale de la Recherche au titre du programme « Investissements d'Avenir » (référence : ANR-10-LABX-0083). Merci à Philippe Martin, Jean-Marie Marandin, Fabián Santiago, Barbara Hemforth, Georges Boulakia, Hiyon Yoo et Elisabeth Delais-Roussarie pour leur aide à différents stades, ainsi qu'à tous les participants.

Références

- ASTESANO C. (2001). *Rythme et accentuation en français : invariance et variabilité stylistique*. Paris : L'Harmattan.
- AVANZI M., LACHERET-DUJOUR A., VICTORRI B. (2008). ANALOR – A tool for semi-automatic annotation of french prosodic structure. Actes de *Speech Prosody 2008*, 119-122.
- BERTRAND R., BLACHE PH., ESPESSE R. FERRE G., MEUNIER C., PRIEGO-VALVERDE B., RAUZY S. (2008). Le CID – Corpus of Interactional Data – Annotation et Exploitation Multimodale de Parole Conversationnelle. *Traitement Automatique des Langues* 49, 1-30.
- DELAIS-ROUSSARIE E., POST B., AVANZI M., ET AL. (2015). Intonational Phonology of French: Developing a ToBI System for French. *Intonation in Romance*. Oxford: Oxford University Press, 63-100.
- DELAIS-ROUSSARIE E., DI CRISTO A. (à paraître). L'accentuation. *La Grande Grammaire du français*. Arles : Actes Sud.
- DE LOOZE C. (2010). *Analyse et Interprétation de l'Empan Temporel des Variations Prosodiques en Français et en Anglais*. Aix-Marseille Université.
- DI CRISTO A. (1999). Le cadre accentuel du français contemporain : essai de modélisation. *Langues* 2, 258-267.
- GODEMENT-BERLINE R. (à paraître). Using a replication task to study prosodic highlighting. Actes de *Speech Prosody 2016*.
- GOLDMAN J.-PH. (2011). EasyAlign: an automatic phonetic alignment tool under Praat. Actes de *Interspeech 2011*, 3233-3236.
- JUN S.-A., FOUGERON C. (2000). A Phonological model of French intonation. *Intonation: Analysis, Modeling and Technology*. Dordrecht : Kluwer Academic Publishers, 209-242.
- LAURENS F., MARANDIN J.-M., PATIN C., YOO H. (2011). The Used and the Possible – The Use of Elicited Conversations in the study of Prosody. Actes de *IDP 2009*, 239-257.
- MARTIN PH. (2009). *Intonation du français*. Paris : Armand Colin.
- MERTENS P. (2004). The Prosogram: Semi-Automatic Transcription of Prosody based on a Tonal Perception Model. Actes de *Speech Prosody 2004*, 2320-2323.
- POST B. (2002). French tonal structures. Actes de *Speech Prosody 2002*, 583-586.
- ROSSI M. (1999). *L'intonation : Le système du français*. Paris : Ophrys.